

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

2003年 9月29日

出 願 番 号

Application Number:

特願2003-337239

パリ条約による外国への出願
に用いる優先権の主張の基礎
となる出願の国コードと出願
番号

The country code and number
of your priority application,
to be used for filing abroad
under the Paris Convention, is

J P 2003-337239

出 願 人

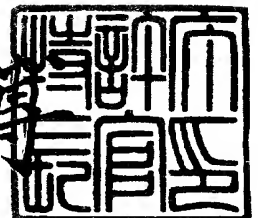
Applicant(s):

株式会社日立製作所

2008年 3月 7日

特許庁長官
Commissioner,
Japan Patent Office

肥塚 雅博



【書類名】	特許願
【整理番号】	340300895
【あて先】	特許庁長官殿
【国際特許分類】	G06F 12/00
【発明者】	
【住所又は居所】	神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I Dシステム事業部内
【氏名】	田村 圭史
【発明者】	
【住所又は居所】	神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I Dシステム事業部内
【氏名】	中山 信一
【発明者】	
【住所又は居所】	神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I Dシステム事業部内
【氏名】	内海 勝広
【特許出願人】	
【識別番号】	000005108
【氏名又は名称】	株式会社日立製作所
【代理人】	
【識別番号】	100095371
【弁理士】	
【氏名又は名称】	上村 輝之
【選任した代理人】	
【識別番号】	100089277
【弁理士】	
【氏名又は名称】	宮川 長夫
【選任した代理人】	
【識別番号】	100104891
【弁理士】	
【氏名又は名称】	中村 猛
【手数料の表示】	
【予納台帳番号】	043557
【納付金額】	21, 000円
【提出物件の目録】	
【物件名】	特許請求の範囲 1
【物件名】	明細書 1
【物件名】	図面 1
【物件名】	要約書 1
【包括委任状番号】	0110323

【書類名】特許請求の範囲

【請求項 1】

第 1 の記憶制御装置と第 2 の記憶制御装置とを通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶システムであって、

前記第 1 の記憶制御装置は、前記上位装置によりアクセスされる少なくとも 1 つ以上の論理ユニットと、この論理ユニットと少なくとも 1 つ以上の記憶デバイスとを接続するようにして設けられる少なくとも 1 つ以上の中間記憶階層とを備えており、

前記中間記憶階層のうち少なくとも 1 つは、前記第 2 の記憶制御装置が有する記憶デバイスに接続されている記憶システム。

【請求項 2】

上位装置及び第 2 の記憶制御装置と通信可能に接続され、前記上位装置からのリクエストに応じたデータ処理を行う記憶制御装置であって、

前記上位装置によりアクセスされる少なくとも 1 つ以上の論理ユニットと、

前記論理ユニットと少なくとも 1 つ以上の記憶デバイスとを接続するようにして設けられる少なくとも 1 つ以上の中間記憶階層とを備え、

前記中間記憶階層のうち少なくとも 1 つは、前記第 2 の記憶制御装置が有する記憶デバイスに接続されている記憶制御装置。

【請求項 3】

前記中間記憶階層は、前記記憶デバイス上に設定される少なくとも 1 つ以上の第 1 の記憶階層と、この第 1 の記憶階層上に設定される少なくとも 1 つ以上の第 2 の記憶階層とを備えて構成されており、前記第 2 の記憶制御装置が有する記憶デバイスは、前記第 1 の記憶階層にマッピングされている請求項 2 に記載の記憶制御装置。

【請求項 4】

前記上位装置からそれぞれ異なる複数の経路を介してアクセス可能な複数の論理ユニットを有し、

前記各論理ユニットは、前記中間記憶階層にそれぞれ接続されている請求項 2 に記載の記憶制御装置。

【請求項 5】

前記第 2 の記憶制御装置が有する前記記憶デバイスへのパス情報を取得するパス情報取得手段を備え、

前記取得したパス情報が複数存在する場合は、前記各パス情報を同一の記憶デバイスへのパス情報として認識する請求項 2 に記載の記憶制御装置。

【請求項 6】

上位装置及び第 2 の記憶制御装置と通信可能に接続され、前記上位装置からのリクエストに応じたデータ処理を行う記憶制御装置の制御方法であって、

前記第 2 の記憶制御装置が有する記憶デバイスへのパス情報を取得するステップと、

前記取得したパス情報を、前記上位装置によりアクセスされる論理ユニットに接続された中間記憶階層にマッピングするステップと、
を含んだことを特徴とする記憶制御装置の制御方法。

【請求項 7】

記憶制御装置に第 2 の記憶制御装置の有する記憶デバイスを内部ボリュームとして設定するためのコンピュータプログラムであって、

前記第 2 の記憶制御装置が有する記憶デバイスへのパス情報を取得するステップと、

前記取得したパス情報を、前記上位装置によりアクセスされる論理ユニットに接続された中間記憶階層にマッピングするステップと、
をコンピュータに実行させるコンピュータプログラム。

【書類名】明細書

【発明の名称】記憶システム及び記憶制御装置

【技術分野】

【0001】

本発明は、記憶システム及び記憶制御装置に関する。

【背景技術】

【0002】

例えば、データセンタ等のような大規模なデータを取り扱うデータベースシステムでは、ホストコンピュータとは別に構成された記憶システムを用いてデータを管理する。この記憶システムは、例えば、ディスクアレイ装置等から構成される。ディスクアレイ装置は、多数の記憶デバイスをアレイ状に配設して構成されるもので、例えば、RAID (Redundant Array of Independent Inexpensive Disks) に基づいて構築されている。記憶デバイス群が提供する物理的な記憶領域上には少なくとも1つ以上の論理ボリューム（論理ユニット）が形成され、この論理ボリュームがホストコンピュータ（より詳しくは、ホストコンピュータ上で稼働するデータベースプログラム）に提供される。ホストコンピュータは、所定のコマンドを送信することにより、論理ボリュームに対してデータの書込み、読み出しを行うことができる。

【0003】

情報化社会の進展等につれて、データベースで管理すべきデータは、日々増大する。このため、より高性能、より大容量の記憶制御装置が求められており、この市場要求に応えるべく、新型の記憶制御装置が開発されている。新型の記憶制御装置を記憶システムを導入する方法としては、2つ考えられる。その一つは、旧型の記憶制御装置と新型の記憶制御装置とを完全に入れ替え、全て新型の記憶制御装置から記憶システムを構成する方法である（特許文献1）。他の一つは、旧型の記憶制御装置からなる記憶システムに新型の記憶制御装置を新たに追加し、新旧の記憶制御装置を併存させる方法である。

【0004】

なお、物理デバイスの記憶領域をセクタ単位で管理し、論理デバイスをセクタ単位で動的に構成する技術も知られている（特許文献2）。

さらに、容量の異なる複数の記憶デバイスから論理デバイスを構築する際に、最も容量の少ない記憶デバイスに合わせてエリアを形成し、残りの容量も最も小さな容量に合わせてエリアを形成するようにした技術も知られている（特許文献3）。

【特許文献1】特表平10-508967号公報

【特許文献2】特開2001-337850号公報

【特許文献3】特開平9-288547号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

旧型の記憶制御装置から新型の記憶制御装置に完全に移行する場合（特許文献1）は、新型の記憶制御装置の機能、性能を利用することができるが、旧型の記憶制御装置を有効に利用することができず、導入コストも増大する。他方、旧型の記憶制御装置と新型の記憶制御装置との併存を図る場合は、記憶システムを構成する記憶制御装置の数が増大し、新旧の記憶制御装置を管理し運用する手間が大きい。

【0006】

また、旧型の記憶制御装置が備える記憶デバイスの応答性が低い場合、この旧型の記憶デバイスを記憶システムに接続することにより、システム全体の性能が低下する。例えば、旧型の記憶デバイスが機械的な動作（ヘッドシーク等）を伴う装置であって、機械的動作時間が長い場合や、旧型の記憶デバイスの備えるデータ転送用バッファの容量が少ない場合等である。

【0007】

さらに、オープン系記憶デバイスとメインフレームや、特定の機能を備えた記憶デバイ

スのみ接続可能なサーバ等のように、旧型の記憶デバイスをそのままでは利用できない場合もある。

【0008】

本発明は、上記の問題点に鑑みてなされたもので、本発明の目的の一つは、例えば、新旧の記憶制御装置のように異なる記憶制御装置を連携させて、記憶資源を有効利用できるようにした記憶システム及び記憶制御装置を提供することにある。

本発明の目的の一つは、旧型の記憶デバイスを新型の記憶デバイスのように利用することが可能な記憶システム及び記憶制御装置を提供することにある。

本発明の目的の一つは、旧型の記憶デバイスの有する利点を活かしつつ、新たな機能を追加できるようにした記憶システム及び記憶制御装置を提供することにある。

本発明の更なる目的は、後述する実施の形態の記載から明らかになるであろう。

【課題を解決するための手段】

【0009】

上記課題を解決すべく、本発明に従う記憶システムは、第1の記憶制御装置と第2の記憶制御装置とを通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶システムであって、第1の記憶制御装置は、上位装置によりアクセスされる少なくとも1つ以上の論理ユニットと、この論理ユニットと少なくとも1つ以上の記憶デバイスとを接続するようにして設けられる少なくとも1つ以上の中間記憶階層とを備えており、中間記憶階層のうち少なくとも1つは、第2の記憶制御装置が有する記憶デバイスに接続されている。

【0010】

ここで、記憶制御装置としては、例えば、ディスクアレイ装置やファイバチャネルスイッチ等を挙げることができる。上位装置としては、例えば、パーソナルコンピュータ、メインフレーム等のコンピュータを挙げることができる。第1の記憶制御装置と第2の記憶制御装置とは、通信ネットワークを介して双方向通信可能に接続されており、第1の記憶制御装置と上位装置も通信ネットワークを介して双方向通信可能に接続されている。また、第2の記憶制御装置と上位装置との間も双方向通信可能に接続することができる。通信ネットワークとしては、例えば、LAN (Local Area Network)、SAN (Storage Area Network)、専用回線、インターネット等を挙げることができる。上位装置からのリクエストに応じたデータ処理としては、例えば、データの読み出し処理、データの書き込み処理等を挙げることができる。第1の記憶制御装置と第2の記憶制御装置とは、同一サイト内に設置されてもよいし、それぞれ異なるサイトに設置されてもよい。また、第1の記憶制御装置が記憶デバイスを備えている場合、この記憶デバイスも中間記憶階層を介して論理ユニットに接続される。

【0011】

第1の記憶制御装置は、少なくとも1つ以上の論理ユニットを備える。この論理ユニットは、例えば、LUN (Logical Unit Number) である。論理ユニットは、上位装置に記憶領域として提供される。上位装置は、論理ユニットへのデータの書き込み、論理ユニットからのデータの読み出し等を行うことができる。論理ユニットの下層には、中間記憶階層が設けられている。この中間記憶階層は、論理ユニットと記憶デバイスとの間を接続するものである。上位装置からのアクセスは、論理ユニットから中間記憶階層を介して記憶デバイスに伝えられる。論理ユニット及び中間記憶階層は、それぞれ複数ずつ設けることができる。少なくとも1つの中間記憶階層は、第2の記憶制御装置の記憶デバイス（外部記憶デバイスまたは第2の記憶デバイスと称することも可能である）に接続されている。

【0012】

つまり、第2の記憶制御装置の記憶デバイスは、中間記憶階層を介して論理ユニットに接続され、論理ユニットを介して上位装置に接続されている。従って、上位装置は、第1の記憶制御装置の記憶デバイスを利用するのと同様にして、何ら意識することなく、第2の記憶制御装置の記憶デバイスを利用することができる。また、例えば、RAID等の機能を中間記憶階層で実現する場合、第2の記憶制御装置の記憶デバイスに第1の記憶制御装置

の有する機能を追加して、上位装置に提供することができる。

【0013】

第2の記憶制御装置の記憶領域を論理ユニットに直接接続することも理論的には可能であるが、中間記憶階層を介さずに直結した場合は、第2の記憶制御装置が有する機能や性能に拘束されてしまい、記憶領域の構成（RAID、分割、拡張等）を変化させたり、機能を追加等することが困難となる。これに対し、第2の記憶制御装置の記憶デバイスを、中間記憶階層を介して論理ユニットに接続することにより、第2の記憶制御装置の機能や性能に依存せずに、新たな記憶資源に再編成して上位装置に提供することができる。例えば、複数の第2の記憶制御装置の各記憶デバイスを1つの中間記憶階層にまとめたり、1つの第2の記憶制御装置の記憶デバイスを複数の中間記憶階層に分割等することができる。

【0014】

本発明の別の観点に従う記憶制御装置は、上位装置及び第2の記憶制御装置と通信可能に接続され、上位装置からのリクエストに応じたデータ処理を行う記憶制御装置であって、上位装置によりアクセスされる少なくとも1つ以上の論理ユニットと、論理ユニットと少なくとも1つ以上の記憶デバイスとを接続するようにして設けられる少なくとも1つ以上の中間記憶階層とを備え、中間記憶階層のうち少なくとも1つは、第2の記憶制御装置が有する記憶デバイスに接続されている。

【0015】

そして、本発明の一態様では、中間記憶階層は、記憶デバイス上に設定される少なくとも1つ以上の第1の記憶階層と、この第1の記憶階層上に設定される少なくとも1つ以上の第2の記憶階層とを備えて構成されており、第2の記憶制御装置が有する記憶デバイスは、第1の記憶階層にマッピングされている。

【0016】

中間記憶階層は、下層に位置する第1の記憶階層と、上層に位置する第2の記憶階層とから二重構造で構成することができ、第1、第2の記憶階層間は接続されている。そして、第1の記憶階層に第2の記憶制御装置の記憶デバイスをマッピングすることにより、第2の記憶制御装置の記憶デバイスを、中間記憶階層を介して論理ユニットに接続し、上位装置に提供することができる。ここで、例えば、第1の記憶階層を識別するための識別情報と、第2の記憶制御装置の記憶デバイスを識別するための情報と、第2の記憶制御装置の記憶デバイスの記憶容量及びデバイス種別（ディスクデバイスかテープデバイスか等）と、第2の記憶制御装置の記憶デバイスのパス情報（WWN（World Wide Name）やLUN）とを対応付けたマッピングテーブルを導入することにより、第2の記憶制御装置の記憶デバイスを第1の記憶階層にマッピングすることができる。

【0017】

本発明の一態様では、上位装置からそれぞれ異なる複数の経路を介してアクセス可能な複数の論理ユニットを有し、各論理ユニットは、中間記憶階層にそれぞれ接続される。

このように、複数の論理ユニットを中間記憶階層に接続することにより、一方の経路が障害等で使用不能な場合でも、他方の経路を介してアクセスすることができ、中間記憶階層に接続された第2の記憶制御装置の記憶デバイスの提供する記憶資源を利用することができる。

【0018】

本発明の一態様では、第2の記憶制御装置が有する記憶デバイスへのパス情報を取得するパス情報取得手段を備え、取得したパス情報が複数存在する場合は、各パス情報を同一の記憶デバイスへのパス情報として認識する。

第2の記憶制御装置の記憶デバイスは、複数の経路を備えている場合がある。例えば、複数のLUNから同一の論理ボリュームにアクセス可能な交代パス構造を有する場合である。この場合、パス情報取得手段は、第2の記憶制御装置の記憶デバイスへの内部的なパス情報、即ち、交代パス構造を意識可能なパス情報を取得する。これにより、第2の記憶制御装置の記憶デバイスが備える交代パス構造も利用することができる。

【0019】

本発明の別の観点に従う制御方法は、上位装置及び第２の記憶制御装置と通信可能に接続され、上位装置からのリクエストに応じたデータ処理を行う記憶制御装置の制御方法であって、第２の記憶制御装置が有する記憶デバイスへのバス情報を取得するステップと、取得したバス情報を、上位装置によりアクセスされる論理ユニットに接続された中間記憶階層にマッピングするステップと、を含んだことを特徴とする。

【００２０】

本発明のさらに別の観点に従うコンピュータプログラムは、記憶制御装置に第２の記憶制御装置の有する記憶デバイスを内部ボリュームとして設定するためのコンピュータプログラムであって、第２の記憶制御装置が有する記憶デバイスへのバス情報を取得するステップと、取得したバス情報を、上位装置によりアクセスされる論理ユニットに接続された中間記憶階層にマッピングするステップと、をコンピュータに実行させる。

【発明を実施するための最良の形態】

【００２１】

以下、図１～図９に基づき、本発明の実施形態を説明する。

本発明では、以下に詳述するように、外部に存在する記憶デバイスを自己の仮想デバイス（VDEV）にマッピングすることにより、外部の記憶デバイスを自己の内部ボリュームとして、上位装置に提供する。

【実施例１】

【００２２】

図１は、本実施例による記憶システムの要部の構成を示すブロック図である。

ホスト装置１０は、例えば、CPU（Central Processing Unit）やメモリ等の情報処理資源を備えたコンピュータ装置であり、例えば、パーソナルコンピュータ、ワークステーション、メインフレーム等として構成される。ホスト装置１０は、例えば、キーボードスイッチやポインティングデバイス、マイクロフォン等の情報入力装置（図示せず）と、例えば、モニタディスプレイやスピーカー等の情報出力装置（図示せず）とを備えている。さらに、ホスト装置１０には、例えば、第１の記憶制御装置２０が提供する記憶領域を使用するデータベースソフトウェア等のアプリケーションプログラム１１と、通信ネットワークCN１を介して第１の記憶制御装置２０にアクセスするためのアダプタ１２とが設けられている。

【００２３】

ホスト装置１０は、通信ネットワークCN１を介して、第１の記憶制御装置２０に接続されている。通信ネットワークCN１としては、例えば、LAN、SAN、インターネット、専用回線、公衆回線等の場合に応じて適宜用いることができる。LANを介するデータ通信は、例えば、TCP/IP（Transmission Control Protocol/Internet Protocol）プロトコルに従って行われる。ホスト装置１０がLANを介して第１の記憶制御装置２０に接続される場合、ホスト装置１０は、ファイル名を指定してファイル単位でのデータ入出力を要求する。一方、ホスト装置１０がSANを介して第１の記憶制御装置２０等に接続される場合、ホスト装置１０は、ファイバチャネルプロトコルに従って、複数のディスク記憶装置（ディスクドライブ）により提供される記憶領域のデータ管理単位であるブロックを単位としてデータ入出力を要求する。通信ネットワークCN１がLANである場合、アダプタ１２は、例えばLAN対応のネットワークカードである。通信ネットワークCN１がSANの場合、アダプタ１２は、例えばホストバスアダプタである。

【００２４】

なお、図中では、ホスト装置１０は、通信ネットワークCN１を介して第１の記憶制御装置２０のみに接続されているが、ホスト装置１０と第２の記憶制御装置４０とを通信ネットワークCN２を介して接続してもよい。第２の通信ネットワークCN２は、例えば、SAN、LAN、インターネット、専用回線、公衆回線等から構成できる。

【００２５】

第１の記憶制御装置２０は、例えば、ディスクアレイサブシステムとして構成されるものである。但し、これに限らず、第１の記憶制御装置２０を、高機能化されたインテリジ

メント型のファイバチャネルスイッチとして構成することもできる。第1の記憶制御装置20は、後述のように、第2の記憶制御装置40の有する記憶資源を自己の論理ボリューム(Logical Unit)としてホスト装置10に提供するものであるから、自己が直接支配するローカルな記憶デバイスを有している必要はない。

【0026】

第1の記憶制御装置20は、コントローラ部と記憶装置部とに大別することができ、コントローラ部は、例えば、複数のチャネルアダプタ(CHA)21と、複数のディスクアダプタ(DKA)22と、コントロールユニット(CU)23と、キャッシュメモリ24と、共有メモリ25と、接続部26とを備えている。

【0027】

各チャネルアダプタ21は、ホスト装置10との間のデータ通信を行うものである。各チャネルアダプタ21は、ホスト装置10と通信を行うための通信ポート21Aを備えている。また、各チャネルアダプタ21は、それぞれCPUやメモリ等を備えたマイクロコンピュータシステムとして構成されており、ホスト装置10から受信した各種コマンドを解釈して実行する。各チャネルアダプタ21には、それぞれを識別するためのネットワークアドレス(例えば、IPアドレスやWWN)が割り当てられており、各チャネルアダプタ21は、それぞれが個別にNAS(Network Attached Storage)として振る舞うことができるようになっていて、複数のホスト装置10が存在する場合、各チャネルアダプタ21は、各ホスト装置10からの要求をそれぞれ個別に受け付けることができる。

【0028】

各ディスクアダプタ22は、記憶装置30の記憶デバイス31、32との間のデータ授受を行うものである。各ディスクアダプタ22は、記憶デバイス31、32に接続するための通信ポート22Aを備えている。また、各ディスクアダプタ22は、CPUやメモリ等を備えたマイクロコンピュータシステムとして構成されている。各ディスクアダプタ22は、チャネルアダプタ21がホスト装置10から受信したデータを、ホスト装置10からのリクエストに基づいて(書込み命令)、所定の記憶デバイス31、32の所定のアドレスに書込み、また、ホスト装置10からのリクエストに基づいて(読み出し命令)、所定の記憶デバイス31、32の所定のアドレスからデータを読み出し、ホスト装置10に送信させる。記憶デバイス31、32との間でデータ入出力を行う場合、各ディスクアダプタ22は、論理的なアドレスを物理的なアドレスに変換する。各ディスクアダプタ22は、記憶デバイス31、32がRAIDに従って管理されている場合は、RAID構成に応じたデータアクセスを行う。

【0029】

コントロールユニット23は、装置全体の作動を制御するものである。コントロールユニット23には、例えば、管理用のコンソール(図示せず)が接続される。コントロールユニット23は、装置内の障害発生を監視してコンソールに表示させたり、コンソールからの指令に基づいて記憶ディスクの閉塞処理等を指示するようになっている。

【0030】

キャッシュメモリ24は、ホスト装置10から受信したデータや、記憶デバイス31、32から読み出したデータを一時的に記憶するものである。共有メモリ25には、制御情報等が格納される。また、共有メモリ25には、ワーク領域が設定されるほか、後述するマッピングテーブルTm等の各種テーブル類も格納される。なお、記憶デバイス31、32のいずれか1つあるいは複数、キャッシュ用のディスクとして使用してもよい。

【0031】

接続部26は、各チャネルアダプタ21、各ディスクアダプタ22、コントロールユニット23、キャッシュメモリ24、共有メモリ25を相互に接続させる。接続部26は、例えば、高速スイッチング動作によってデータ伝送を行う超高速クロスバススイッチ等のような高速バスとして構成することができる。

【0032】

記憶装置30は、複数の記憶デバイス31を備えている。記憶デバイス31としては、

例えば、ハードディスク、フレキシブルディスク、磁気テープ、半導体メモリ、光ディスク等のようなデバイスを用いることができる。記憶装置30内に点線で示される記憶デバイス32は、第2の記憶制御装置40の有する記憶デバイス42を第1の記憶制御装置20側に取り込んだ状態を示すものである。即ち、本実施例では、第1の記憶制御装置20から見て外部に存在する記憶デバイス42を、第1の記憶制御装置20の内部記憶デバイスとして認識し、ホスト装置10に外部記憶デバイス42の記憶資源を提供する。

【0033】

第2の記憶制御装置40は、通信ポート41と記憶デバイス42とを備えている。このほか、チャネルアダプタやディスクアダプタ等を備えることもできるが、第2の記憶制御装置40の詳細な構造は、本発明の要旨ではないので説明を割愛する。第2の記憶制御装置40は、通信ネットワークCN2を介して第1の記憶制御装置20に接続されており、第2の記憶制御装置40の記憶デバイス42は、第1の記憶制御装置20の内部記憶デバイスとして扱われるようになっている。

【0034】

図2を参照する。図2は、第1の記憶制御装置20及び記憶デバイス32の1つの論理的な概略構造を示す模式図である。図に示すように、第1の記憶制御装置20は、下層側から順番に、VDEV101と、LDEV102と、LUN103とからなる3層の記憶階層を有している。

【0035】

VDEV101は、論理的な記憶階層の最下位に位置する仮想デバイス（Virtual Device）である。VDEV101は、物理的な記憶資源を仮想化したものであり、RAID構成を適用することができる。即ち、1つの記憶デバイス31から複数のVDEV101を形成することもできるし（スライシング）、複数の記憶デバイス31から1つのVDEV101を形成することもできる（ストライピング）。図2中の左側に示すVDEV101は、例えば、所定のRAID構成に従って記憶デバイス31を仮想化している。

【0036】

一方、図2中の右側に示すVDEV101は、第2の記憶制御装置40の記憶デバイス42をマッピングすることにより構成されている。即ち、本実施例では、第2の記憶制御装置40の記憶デバイス42により提供される論理ボリューム（LDEV）を、後述のマッピングテーブルTmを用いてVDEV101にマッピングすることにより、第1の記憶制御装置20の内部ボリュームとして使用できるようになっている。図に示す例では、4つの記憶デバイス42A～42Dをストライピングすることにより、VDEV101を構築している。各記憶デバイス42A～42Dには、それぞれの通信ポート41A～41DからそれぞれのLUN43A～43Dを特定することにより、それぞれ個別にアクセスすることができる。各通信ポート41A～41Dには、ユニークな識別情報であるWWNが割り当てられており、また、各LUN43A～43Dには、LUN番号が設定されているので、WWN及びLUN番号の組合せによって記憶デバイスを特定できる。

【0037】

VDEV101の上には、LDEV102が設けられている。LDEV102は、仮想デバイス（VDEV）を仮想化した論理デバイス（論理ボリューム）である。1つのVDEV101から2つのLDEV102に接続することもできるし、複数のVDEV101から1つのLDEV102に接続することもできる。LDEV102には、それぞれのLUN103を介してアクセスすることができる。このように、本実施例では、LUN103と記憶デバイス42との間に位置する中間記憶階層（VDEV101、LDEV102）に記憶デバイス42を接続することにより、外部の記憶デバイス42を第1の記憶制御装置20の内部ボリュームの1つとして利用できるようにしている。

【0038】

図3は、他の論理的概略構造を示す模式図である。図3では、第2の記憶制御装置40の記憶デバイス42により提供されるLDEV50は、複数の経路を備えた交代パス構成を有している。即ち、各記憶デバイス42の上には、論理ボリュームであるLDEV50

が構築されており、このLDEV50には、2つの経路（アクセスデータパス）を介してそれぞれアクセスすることができる。一方の経路は、第1の通信ポート41（1）からLUN43を介してLDEV50に到達し、他方の経路は、第2の通信ポート41（2）から別のLUN43を介してLDEV50に到達する。従って、仮にいずれか一方の経路が障害等で使用不能な場合でも、他方の経路を介してLDEV50にアクセスすることができる。複数の経路からそれぞれアクセス可能な場合、一方の経路からデータを利用中に、他方の経路からアクセスしてデータを更新等することがないように、必要なデータ保護等が行われる。

【0039】

なお、図3に示す例では、第1の記憶制御装置20は、第2の記憶制御装置40の記憶資源（LDEV50）を自己のVDEV101にマッピングすることにより、外部のLDEV50を内部のLDEV102として利用している。また、複数のLDEV102を1つのVDEV101上に設定し、このVDEV101には外部のLDEV50が複数の経路を介してマッピングされている。ホスト装置10は、LUN103のみを認識しており（結果的にLDEV102まで認識しており）、LUN103よりも下の構造は、ホスト装置10に対して隠されている。複数のLDEV102は、それぞれ同一のVDEV101を利用し、このVDEV101は複数の経路を介して同一のLDEV50に接続されている。従って、図3に示す例では、第2の記憶制御装置40の有する交代パス構造を利用して、第1の記憶制御装置20の冗長性を高めることができる。

【0040】

次に、図4を参照する。図4は、外部の記憶デバイス42（詳しくは、外部の記憶デバイス42により提供されるLDEV50）を、VDEV101にマッピングするためのテーブル構造の一例を示す。

【0041】

マッピングテーブルT_mは、例えば、VDEV101をそれぞれ識別するためのVDEV番号と外部の記憶デバイス42の情報とをそれぞれ対応付けることにより、構成することができる。外部デバイス情報としては、例えば、デバイス識別情報と、記憶デバイス42の記憶容量と、デバイスの種別を示す情報と（例えば、テープ系デバイスかディスク系デバイスか等）、記憶デバイス42へのパス情報とを含んで構成することができる。また、パス情報は、各通信ポート41に固有の識別情報（WWN）と、LUN43を識別するためのLUN番号とを含んで構成できる。なお、図4中に示すデバイス識別情報やWWN等は、説明の便宜上の値であって特に意味はない。また、図4中の下側に示すVDEV番号「3」のVDEV101には、3個のパス情報が対応付けられている。即ち、このVDEV101（#3）にマッピングされる外部記憶デバイス42は、その内部に3つの経路を有する交代パス構造を備えているが、VDEV101（#3）には、この交代パス構造を認識してマッピングされている。これら3つの経路のいずれを通っても同一の記憶領域にアクセスできることが判明しているため、いずれか1つまたは2つの経路に障害等が発生した場合でも、残りの正常な経路を介して所望のデータにアクセスできる。

図4に示すようなマッピングテーブルT_mを採用することにより、第1の記憶制御装置20内の1つ以上のVDEV101に対し、1つまたは複数の外部の記憶デバイス42をマッピングすることができる。

【0042】

次に、図5を参照して、外部の記憶デバイス42をVDEV101にマッピングする方法の一例を説明する。図5は、マッピング時に、第1の記憶制御装置20と第2の記憶制御装置40との間で行われる処理の要部を示すタイムチャートである。

【0043】

まず、第1の記憶制御装置20は、チャネルアダプタ21のイニシエータポート（21A）を介して、第2の記憶制御装置40にログインする（S1）。第2の記憶制御装置40が、第1の記憶制御装置20のログインに対して応答を返すことにより、ログインが完了する（S2）。次に、第1の記憶制御装置20は、例えば、SCSI（Small Computer Sys

tem Interface) 規格で定められている照会コマンド (inquiryコマンド) を、第2の記憶制御装置40に送信し、第2の記憶制御装置40の有する記憶デバイス42の詳細について応答を求める (S3)。

【0044】

照会コマンドは、照会先の装置の種類及び構成を明らかにするために用いられるもので、照会先装置の有する階層を透過してその物理的構造を把握することができる。照会コマンドを使用することにより、第1の記憶制御装置20は、例えば、装置名、デバイスタイプ、製造番号 (プロダクトID)、LDEV番号、各種バージョン情報、ベンダID等の情報を第2の記憶制御装置40から取得できる (S4)。第2の記憶制御装置40は、問合せされた情報を第1の記憶制御装置20に送信し、応答する (S5)。

【0045】

第1の記憶制御装置20は、第2の記憶制御装置40から取得した情報を、マッピングテーブルTmの所定箇所に登録する (S6)。次に、第1の記憶制御装置20は、第2の記憶制御装置40から記憶デバイス42の記憶容量を読み出す (S7)。第2の記憶制御装置40は、第1の記憶制御装置20からの問合せに対して、記憶デバイス42の記憶容量を返信し (S8)、応答を返す (S9)。第1の記憶制御装置20は、記憶デバイス42の記憶容量をマッピングテーブルTmの所定箇所に登録する (S10)。

【0046】

以上の処理を行うことにより、マッピングテーブルTmを構築できる。第1の記憶制御装置20のVDEV101にマッピングされた外部の記憶デバイス42 (外部LUN、即ち外部のLDEV50) との間でデータの入出力を行う場合は、後述する他のテーブルを参照してアドレス変換等を行う。

【0047】

図6～図8を参照して、第1の記憶制御装置20と第2の記憶制御装置40との間のデータ入出力について説明する。まず最初に、データを書き込む場合について、図6及び図7に基づいて説明する。図6は、データ書き込み時の処理を示す模式図である。図7は、図6中の処理の流れを各種テーブルとの関係で示す説明図である。

【0048】

ホスト装置10は、第1の記憶制御装置20が提供する論理ボリューム (LDEV102) にデータを書き込むことができる。例えば、SANの中に仮想的なSANサブネットを設定するゾーニングや、アクセス可能なLUNのリストをホスト装置10が保持するLUNマスキングという手法により、ホスト装置10を特定のLDEV102に対してのみアクセスさせるように設定できる。

【0049】

ホスト装置10がデータを書き込もうとするLDEV102が、VDEV101を介して内部の記憶デバイスである記憶デバイス31に接続されている場合、通常の処理によってデータが書き込まれる。即ち、ホスト装置10からのデータは、いったんキャッシュメモリ24に格納され、キャッシュメモリ24からディスクアダプタ22を介して、所定の記憶デバイス31の所定アドレスに格納される。この際、ディスクアダプタ22は、論理的なアドレスを物理的なアドレスに変換する。また、RAID構成の場合、同一のデータが複数の記憶デバイス31に記憶等される。

【0050】

これに対し、ホスト装置10が書き込もうとするLDEV102が、VDEV102を介して外部の記憶デバイス42に接続されている場合、図6に示すような流れでデータが書き込まれる。図6 (a) は記憶階層を中心に示す流れ図であり、図6 (b) はキャッシュメモリ24の使われ方を中心に示す流れ図である。

【0051】

ホスト装置10は、書き込み先のLDEV102を特定するLDEV番号とこのLDEV102にアクセスするための通信ポート21Aを特定するWWNとを明示して、書き込みコマンド (Write) を発行する (S21)。第1の記憶制御装置20は、ホスト装置10か

らの書込みコマンドを受信すると、第2の記憶制御装置40に送信するための書込みコマンドを生成し、第2の記憶制御装置40に送信する(S22)。第1の記憶制御装置20は、ホスト装置10から受信した書込みコマンド中の書込み先アドレス情報等を、外部LDEV50に合わせて変更することにより、新たな書込みコマンドを生成する。

【0052】

次に、ホスト装置10は、書き込むべきデータを第1の記憶制御装置40に送信する(S23)。第1の記憶制御装置20に受信されたデータは、LDEV102からVDEV101を介して(S24)、外部のLDEV50に転送される(S26)。ここで、第1の記憶制御装置20は、ホスト装置10からのデータをキャッシュメモリ24に格納した時点で、ホスト装置10に対し書込み完了の応答(Good)を返す(S25)。第2の記憶制御装置40は、第1の記憶制御装置20からデータを受信した時点で(あるいは記憶デバイス42に書込みを終えた時点で)、書込み完了報告を第1の記憶制御装置20に送信する(S26)。即ち、第1の記憶制御装置20がホスト装置10に対して書込み完了を報告する時期(S25)と、実際にデータが記憶デバイス42に記憶される時期とは相違する(非同期方式)。従って、ホスト装置10は、実際にデータが記憶デバイス42に格納される前にデータ書込み処理から解放され、別の処理を行うことができる。

【0053】

図6(b)を参照する。キャッシュメモリ24には、多数のサブブロック24Aが設けられている。第1の記憶制御装置20は、ホスト装置10から指定された論理ブロックアドレスをサブブロックのアドレスに変換し、キャッシュメモリ24の所定箇所にデータを格納する(S24)。

【0054】

図7を参照して、各種テーブルを利用してデータが変換される様子を説明する。図7の上部に示すように、ホスト装置10は、所定の通信ポート21Aに対し、LUN番号(LUN#)及び論理ブロックアドレス(LBA)を指定してデータを送信する。第1の記憶制御装置20は、LDEV102用に入力されたデータ(LUN#+LBA)を、図7(a)に示す第1の変換テーブルT1に基づいて、VDEV101用のデータに変換する。第1の変換テーブルT1は、内部のLUN103を指定するデータをVDEV101用データに変換するための、LUN-LDEV-VDEV変換テーブルである。このテーブルT1は、例えば、LUN番号(LUN#)と、そのLUN103に対応するLDEV102の番号(LDEV#)及び最大スロット数と、LDEV102に対応するVDEV101の番号(VDEV#)及び最大スロット数等を対応付けることにより構成される。このテーブルT1を参照することにより、ホスト装置10からのデータ(LUN#+LBA)は、VDEV101用のデータ(VDEV#+SLOT#+SUBBLOCK#)に変換される。

【0055】

次に、第1の記憶制御装置20は、図7(b)に示す第2の変換テーブルT2を参照して、VDEV101用のデータを、第2の記憶制御装置40の外部LUN(LDEV50)用に送信して記憶させるためのデータに変換する。第2の変換テーブルT2には、例えば、VDEV101の番号(VDEV#)と、そのVDEV101からのデータを第2の記憶制御装置40に送信するためのイニシエータポートの番号と、データ転送先の通信ポート41を特定するためのWWNと、その通信ポートを介してアクセス可能なLUN番号とが対応付けられている。この第2の変換テーブルT2に基づいて、第1の記憶制御装置20は、記憶させるべきデータの宛先情報を、イニシエータポート番号#+WWN+LUN#+LBAの形式に変換する。このように宛先情報が変更されたデータは、指定されたイニシエータポートから通信ネットワークCN2を介して、指定された通信ポート41に到達する。そして、データは、指定されたLUN43でアクセス可能なLDEV50の所定の場所に格納される。LDEV50は、複数の記憶デバイス42上に仮想的に構築されているので、データのアドレスは物理アドレスに変換されて、所定のディスクの所定アドレスに格納される。

【0056】

図7(c)は、別の第2の変換テーブルT2aを示す。この変換テーブルT2aは、外部記憶デバイス42に由来するVDEV101に、ストライプやRAIDを適用する場合に使用される。変換テーブルT2aは、VDEV番号(VDEV#)と、ストライプサイズと、RAIDレベルと、第2の記憶制御装置40を識別するための番号(SS#(ストレージシステム番号))と、イニシエータポート番号と、通信ポート41のWWN及びLUN43の番号とを対応付けることにより構成されている。図7(c)に示す例では、1つのVDEV101は、SS#(1, 4, 6, 7)で特定される合計4つの外部記憶制御装置を利用してRAID1を構成する。また、SS#1に割り当てられている3個のLUN(#0, #0, #4)は、同一デバイス(LDEV#)に設定されている。なお、LUN#0のボリュームは、2個のアクセスデータパスを有する交代パス構造を備えている。このように、本実施例では、外部に存在する複数の論理ボリューム(LDEV)からVDEV101を構成することにより、ストライピングやRAID等の機能を追加した上でホスト装置10に提供することができる。

【0057】

図8を参照して、第2の記憶制御装置40のLDEV50からデータを読み出す場合の流れを説明する。

まず、ホスト装置10は、通信ポート21Aを指定して第1の記憶制御装置20にデータの読み出しコマンドを送信する(S31)。第1の記憶制御装置20は、読み出しコマンドを受信すると、要求されたデータを第2の記憶制御装置40から読み出すべく、読み出しコマンドを生成する。第1の記憶制御装置20は、生成した読み出しコマンドを第2の記憶制御装置40に送信する(S32)。第2の記憶制御装置40は、第1の記憶制御装置20から受信した読み出しコマンドに応じて、要求されたデータを記憶デバイス42から読み出して、第1の記憶制御装置20に送信し(S33)、正常に読み出しが完了した旨を報告する(S35)。第1の記憶制御装置20は、図8(b)に示すように、第2の記憶制御装置40から受信したデータを、キャッシュメモリ24の所定の場所に格納させる(S34)。

【0058】

第1の記憶制御装置20は、キャッシュメモリ24に格納されたデータを読み出し、アドレス変換を行った後、LUN103等を介してホスト装置10にデータを送信し(S36)、読み出し完了報告を行う(S37)。これらデータ読み出し時の一連の処理では、図7と共に述べた変換操作が逆向きで行われる。

【0059】

図8では、ホスト装置10からの要求に応じて、第2の記憶制御装置40からデータを読み出し、キャッシュメモリ24に保存するかのように示している。しかし、これに限らず、外部のLDEV50に記憶されているデータの全部または一部を、予めキャッシュメモリ24に記憶させておくこともできる。この場合、ホスト装置10からの読み出しコマンドに対し、直ちにキャッシュメモリ24からデータを読み出してホスト装置10に送信することができる。

【0060】

以上詳述した通り、本実施例によれば、外部の記憶デバイス42(正確には、外部のLDEV50)をVDEV101にマッピングする構成のため、外部の論理ボリュームを内部の論理ボリュームのように、あるいは、外部の記憶デバイス42を仮想的な内部記憶デバイスのように、取り扱うことができる。従って、第2の記憶制御装置40が、ホスト装置10と直接接続することができない旧型の装置である場合でも、新型の第1の記憶制御装置20が介在することにより、旧型装置の記憶資源を第1の記憶制御装置20の記憶資源として再利用し、ホスト装置10に提供することができる。これにより、旧型の記憶制御装置を新型の記憶制御装置20に統合して記憶資源を有効に利用できる。

【0061】

また、第1の記憶制御装置20が高性能・高機能の新型装置の場合、第1の記憶制御装

置 20 の有する高性能なコンピュータ資源（キャッシュ容量や CPU 処理速度等）によって第 2 の記憶制御装置 40 の低性能を隠すことができ、外部の記憶デバイス 42 を活用した仮想的な内部ボリュームを用いて、高性能なサービスをホスト装置 10 に提供することができる。

【0062】

さらに、外部の記憶デバイス 42 上に構築される LDEV 50 に、例えば、ストライピング、拡張、分割、RAID 等の機能を追加して使用することができる。従って、外部のボリュームを直接 LUN 103 にマッピングする場合に比較して、利用の自由度が高まり、使い勝手が向上する。

【0063】

また、外部の論理ボリュームを内部の論理ボリュームのように使用できるため、通常の内部ボリュームである LDEV 102 に対して第 1 の記憶制御装置 20 が利用可能な諸機能を、仮想的な内部ボリューム（LDEV 50 に接続される LDEV）に対しても適用できる。利用可能な諸機能としては、例えば、MRCF、リモートコピー、CVS、LUSE 等を挙げることができる。ここで、MRCF（Multiple RAID Coupling Feature）とは、ホスト装置 10 をデータが経路することなく（ホストフリー）、論理ボリュームのレプリカを作成可能な機能である。リモートコピーとは、ローカルサイトに設置されたプライマリボリュームとリモートサイトに設置されたセカンダリボリュームとの記憶内容を同期させる機能である。CVS（Customizable Volume Size）とは、論理ボリュームのサイズを標準的なサイズ以上の任意のサイズで設定可能な可変ボリューム機能である。LUSE（LU Size Expansion）とは、複数の論理ボリュームを 1 つの論理ボリュームに統合し、ホスト装置 10 から認識可能な LUN 数を少なくさせる LUN サイズ拡張機能である。

【0064】

さらに、外部の論理ボリュームから構築される VDEV 101 を複数の LDEV 102 にそれぞれ接続することができるため、ホスト装置 10 を各 LDEV 102 の LUN 103 にそれぞれ接続することにより、交代パス構造を得ることができ、また、負荷分散効果も得ることができる。

【0065】

また、照会コマンドにより第 2 の記憶制御装置 40 の有する交代パス構造まで把握して VDEV 101 にマッピングするため、第 2 の記憶制御装置 40 の有する交代パス構造も継承することができ、記憶システムの冗長性を高めることができる。

【0066】

これに対し、背景技術の欄で挙げた従来技術（特許文献 2）は、ローカルな記憶デバイス、即ち、記憶制御装置が直接支配下に置いている記憶デバイスのボリュームをセクタ単位で再構成するだけのものであり、本実施例のように、外部の記憶デバイスを仮想的な内部の記憶デバイスとして取り扱うものではない。また、他の従来技術（特許文献 3）は、ローカルな記憶デバイスの記憶容量に合わせてエリアのサイズを最適化するもので、本実施例のように、外部の記憶デバイス 42 を仮想的な内部記憶デバイスとして利用可能な技術ではない。

【実施例 2】

【0067】

次に、図 9 を参照して本発明の第 2 実施例を説明する。本実施例の特徴は、第 2 の記憶制御装置 40 の有する交代パス構造を検証する点にある。

図 9 は、第 1 の記憶制御装置 20 により実行される交代パス構造の検証処理概要を示すフローチャートである。第 1 の記憶制御装置 20 は、マッピングテーブル T_m を参照し、交代パスとして認識された 1 セット（通常は 2 個）のアクセスデータパスを選択する（S41）。

【0068】

次に、第 1 の記憶制御装置 20 は、選択された各パスを介して、所定のアドレスからそれぞれデータを読み出し（S42）、各パスからそれぞれ読み出されたデータが一致する

可否かを判定する（S43）。両方のパスからそれぞれ読み出されたデータが一致する場合は（S43:YES）、交代パス構造が構築されていると一応判断することができる。しかし、各パスがそれぞれ異なるボリュームに接続されており、それぞれの読み出し先アドレスに偶然同一のデータが格納されている可能性も残されている。

【0069】

そこで、本実施例では、第2段階の確認として、一方のパスを介して所定のアドレスに特定のデータを書き込む（S44）。この書き込みデータは、S42で読み出したデータと異なっている。次に、他方のパスを介して所定のアドレスから再びデータを読み出し、この読み出されたデータが、S44で書き込まれた特定のデータと一致するか否かを判断する（S45）。両データが一致する場合は（S45:YES）、交代パス構造が構築されていることが検証される（S46）。

【0070】

第1段階のテスト（S43）または第2段階のテスト（S45）のいずれかで「NO」と判定された場合は、マッピングテーブルTmに登録された交代パス構造が誤っているものと判断し、エラー処理を行う（S47）。このエラー処理としては、マッピングテーブルTmの再構築等を挙げることができる。

【0071】

このように、照会コマンドにより把握された第2の記憶制御装置40の交代パス構造を検証する処理を備えるため、信頼性を向上させることができる。また、各パスからそれぞれ読み出したデータが一致するか否かを判定する第1段階のテストと、読み出したデータと異なるデータを書き込んで、再び各パスからデータをそれぞれ読み出して一致するか否かを判定する第2段階のテストとの2段階で検証を行うため、検証の信頼性を高めることができる。

【0072】

図9に示す処理は、例えば、以下のように表現することができる。

表現1．複数のパスを選択するステップ（S41）と、これら選択された各パスからそれぞれデータを読み出すステップ（S42）と、これら読み出された各データが一致するか否かを判定するステップ（S43）とを備えた交代パス構造の検証方法（または検証方法をコンピュータに実行させるプログラム）。

表現2．複数のパスを選択するステップ（S41）と、これら選択された各パスからそれぞれデータを読み出すステップ（S42）と、これら読み出された各データが一致するか否かを判定するステップ（S43）と、読み出された各データが一致する場合（S43:YES）は、前記選択された各パスのうちいずれか1つのパスから前記読み出されたデータとは異なる別のデータを書き込むステップ（S44）と、前記選択された各パスから（前記別のデータを書き込むパスを含めてもよいし、除いてもよい）再度データを読み出すステップ（S45）と、これら読み出されたデータが前記別のデータと一致するか否かを判定するステップ（S45）と、を備えた交代パス構造の検証方法（または検証方法をコンピュータに実行させるプログラム）。

表現3．外部の記憶制御装置の有する外部記憶デバイスを自己の有する内部記憶デバイスであるかのように上位装置に提供する記憶制御装置において交代パス構造を検証するための検証方法であって、前記外部記憶デバイスへ繋がる複数のパスを選択するステップと、これら選択された各パスを介して前記外部記憶デバイスからそれぞれデータを読み出すステップと、これら読み出された各データが一致するか否かを判定するステップと、読み出された各データが一致する場合は、前記選択された各パスのうちいずれか1つのパスから前記読み出されたデータとは異なる別のデータを書き込むステップと、前記選択された各パスを介して前記外部記憶デバイスから再度データを読み出すステップと、これら読み出されたデータが前記別のデータと一致するか否かを判定するステップと、を備えた交代パス構造の検証方法。

【0073】

なお、本発明は、上述した各実施例に限定されない。当業者であれば、本発明の範囲内

で、種々の追加や変更等を行うことができる。前記各実施例では、ディスクアレイ装置を中心に説明したが、これに限らず、インテリジェント化されたファイバチャネルスイッチにも適用することができる。

【図面の簡単な説明】

【0074】

【図1】 本発明の実施例に係わる記憶システムの全体構成を示すブロック図である。

【図2】 記憶システムの論理的構成の概要を示す模式図である。

【図3】 他の論理的構成の概要を示す模式図である。

【図4】 マッピングテーブルの概要を示す説明図である。

【図5】 マッピングテーブルを構築するための処理の流れを示す説明図である。

【図6】 内部ボリュームとして仮想化される外部の記憶デバイスにデータを書き込む場合の概念図である。

【図7】 書込みデータのアドレス変換の様子を模式的に示す説明図である。

【図8】 内部ボリュームとして仮想化される外部の記憶デバイスからデータを読み出す場合の概念図である。

【図9】 本発明の他の実施例に係る交代パス構造を検証するための処理を示すフローチャートである。

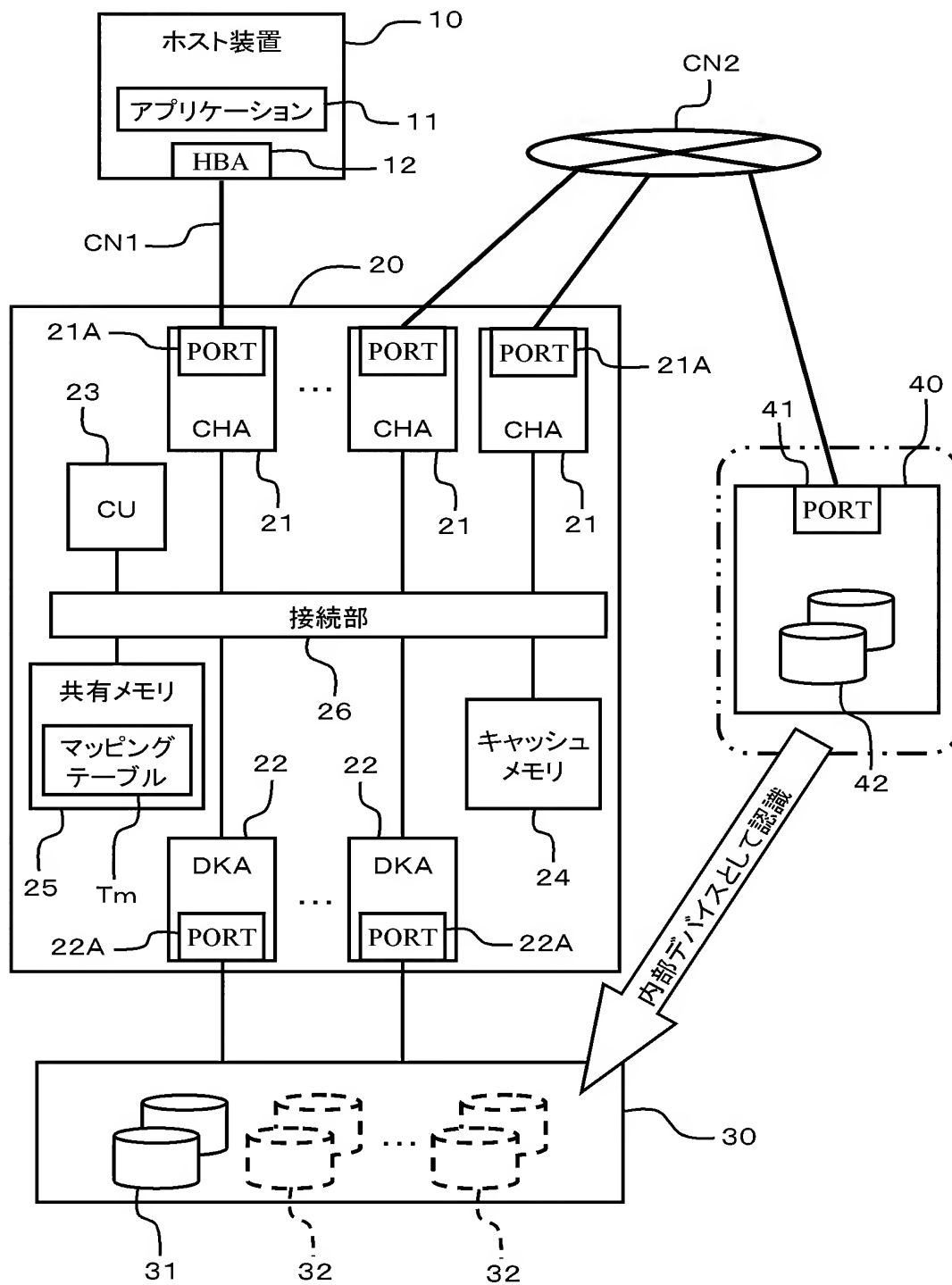
【符号の説明】

【0075】

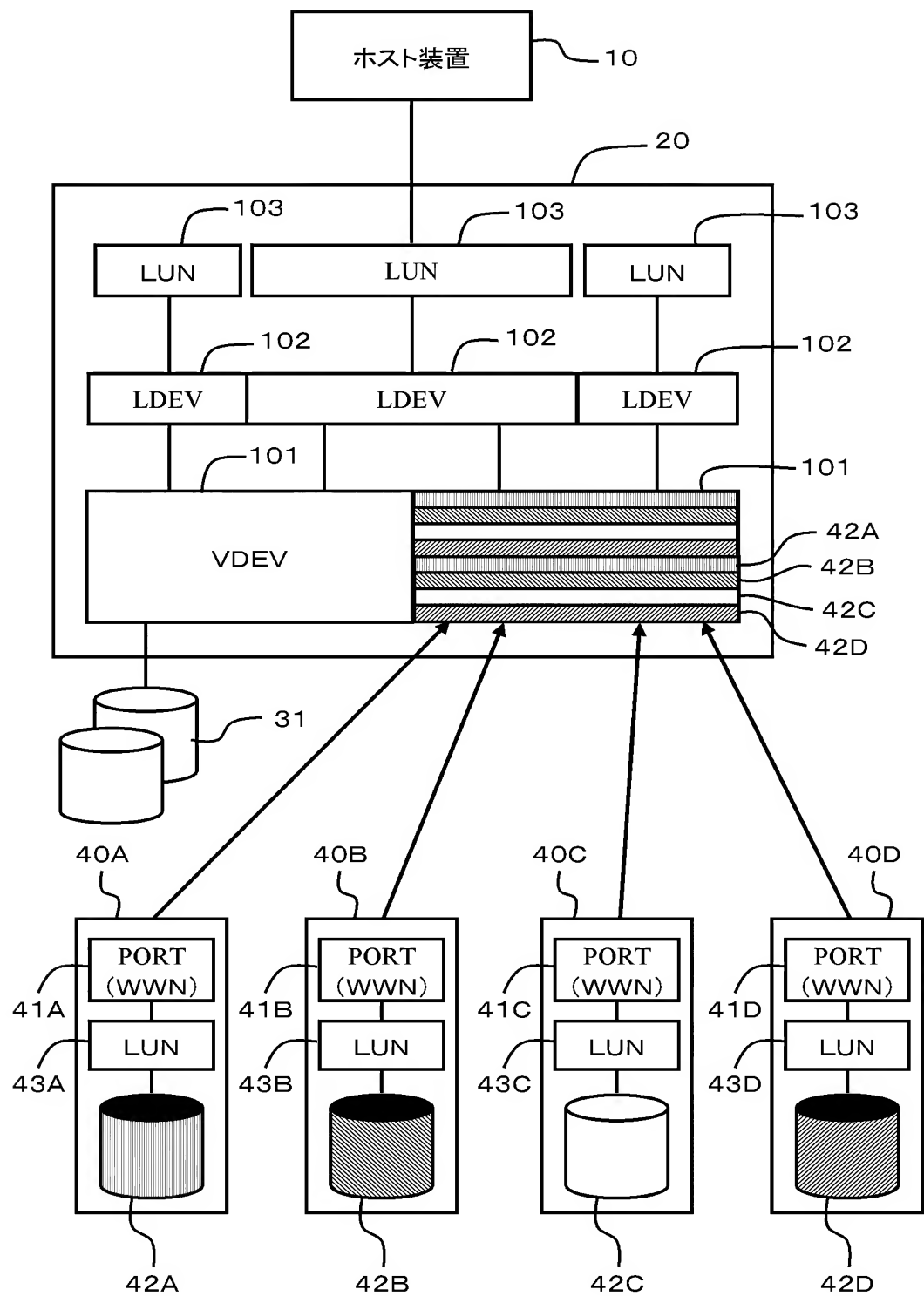
10…ホスト装置、11…アプリケーションプログラム、12…アダプタ、20…第1の記憶制御装置、21…チャネルアダプタ、21A…通信ポート、22…ディスクアダプタ、22A…通信ポート、23…コントロールユニット、24…キャッシュメモリ、24A…サブブロック、25…共有メモリ、26…接続部、30…記憶装置、31…記憶デバイス、32…記憶デバイス（仮想化された内部記憶デバイス）、40…第2の記憶制御装置、41…各通信ポート、42…各記憶デバイス、CN1…第1の通信ネットワーク、CN2…第2の通信ネットワーク、50…LDEV（仮想化された内部論理デバイス）、101…VDEV（仮想デバイス）、102…LDEV（論理デバイス）、103…LUN、T1…第1の変換テーブル、T2…第2の変換テーブル、T2a…別の第2の変換テーブル、Tm…マッピングテーブル

【書類名】 図面

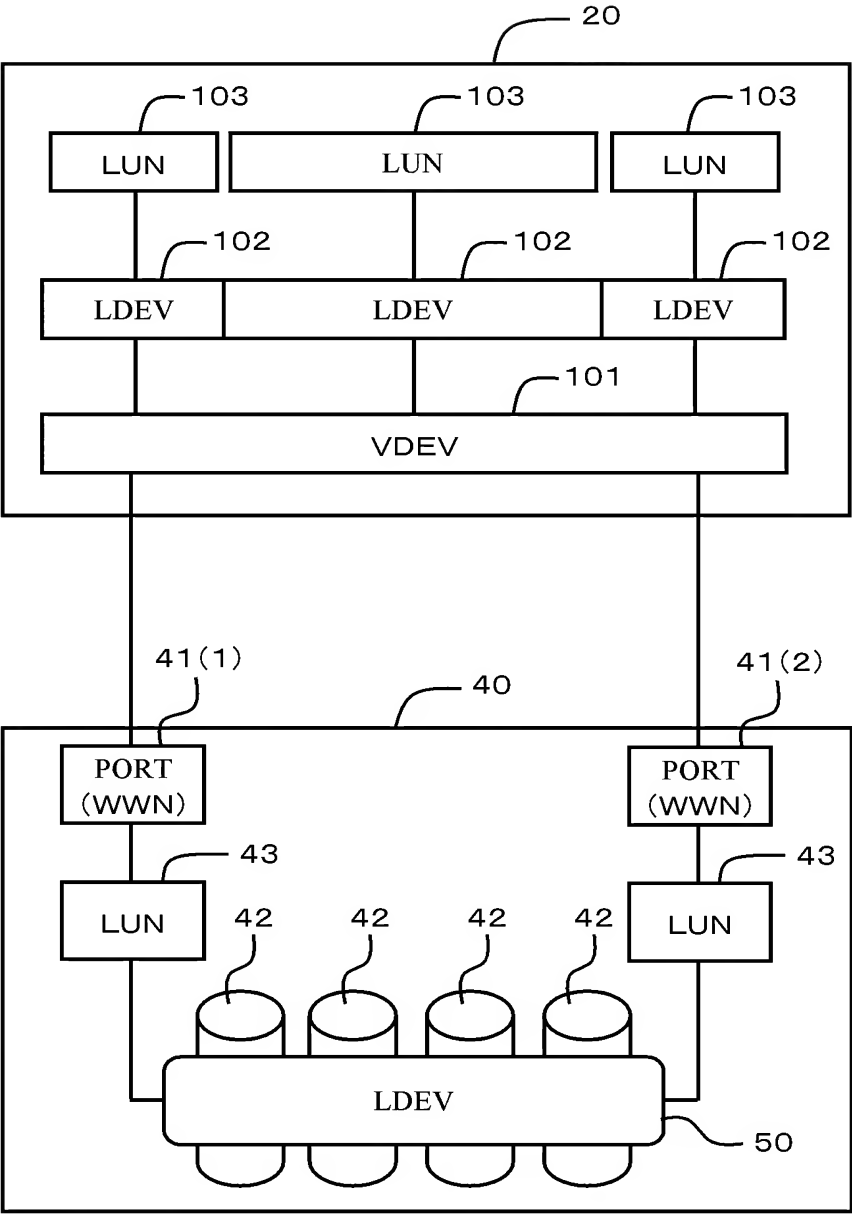
【図 1】



【図2】



【図3】



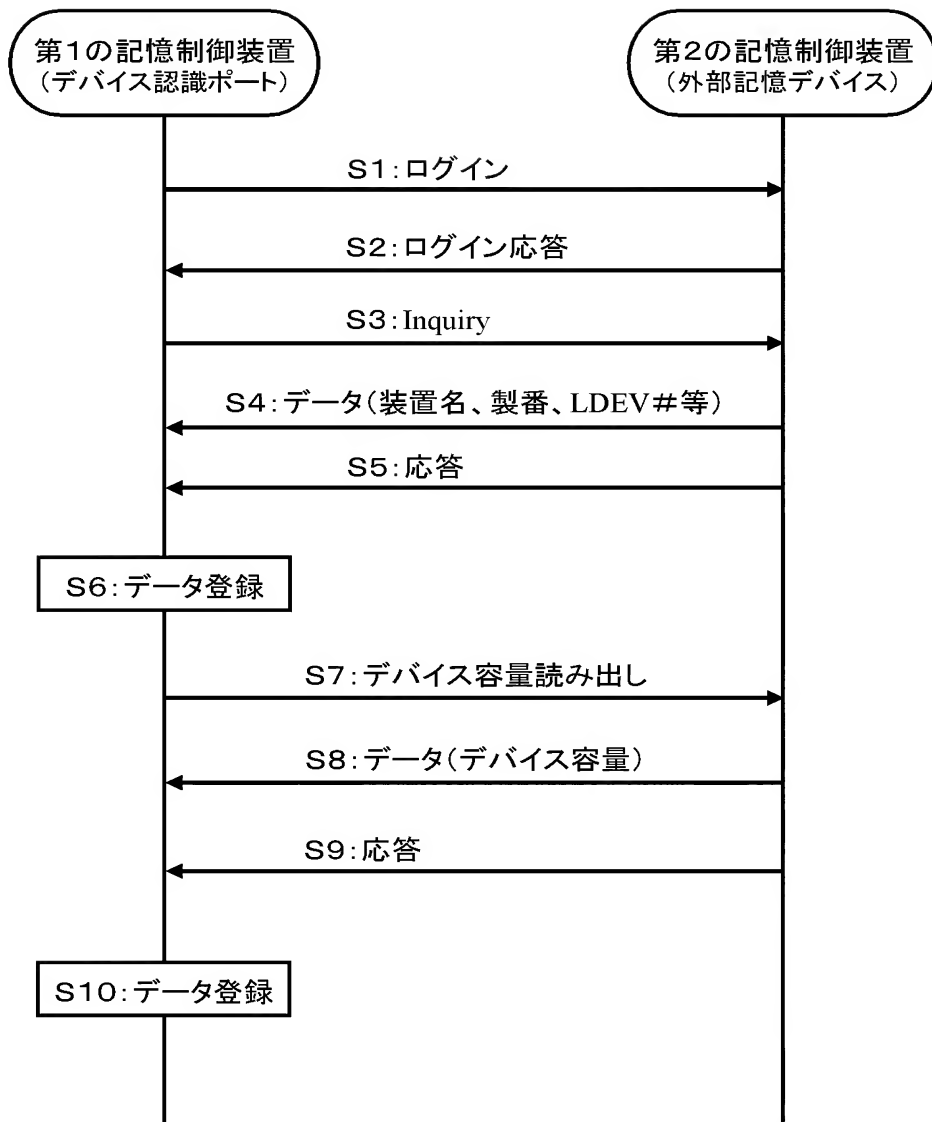
【図 4】

VDEV	外部デバイス情報			
	デバイス 識別情報	容量(KB)	デバイス 種別	パス情報 WWN LUN
0	DRFGTFNEIEK	657,456	DISK	0xAABBCCDD 0
1	ADRFGTFNEIE	89,854	DISK	0xAABBEFFF 3
2	GGRRFFDDERT	—	TAPE	0x445566AAB 5
3	AABBCDDDEE	5,544,223	DISK	0x77DE12345 6
				0x77DE12345 3
				0x377DE7890 5

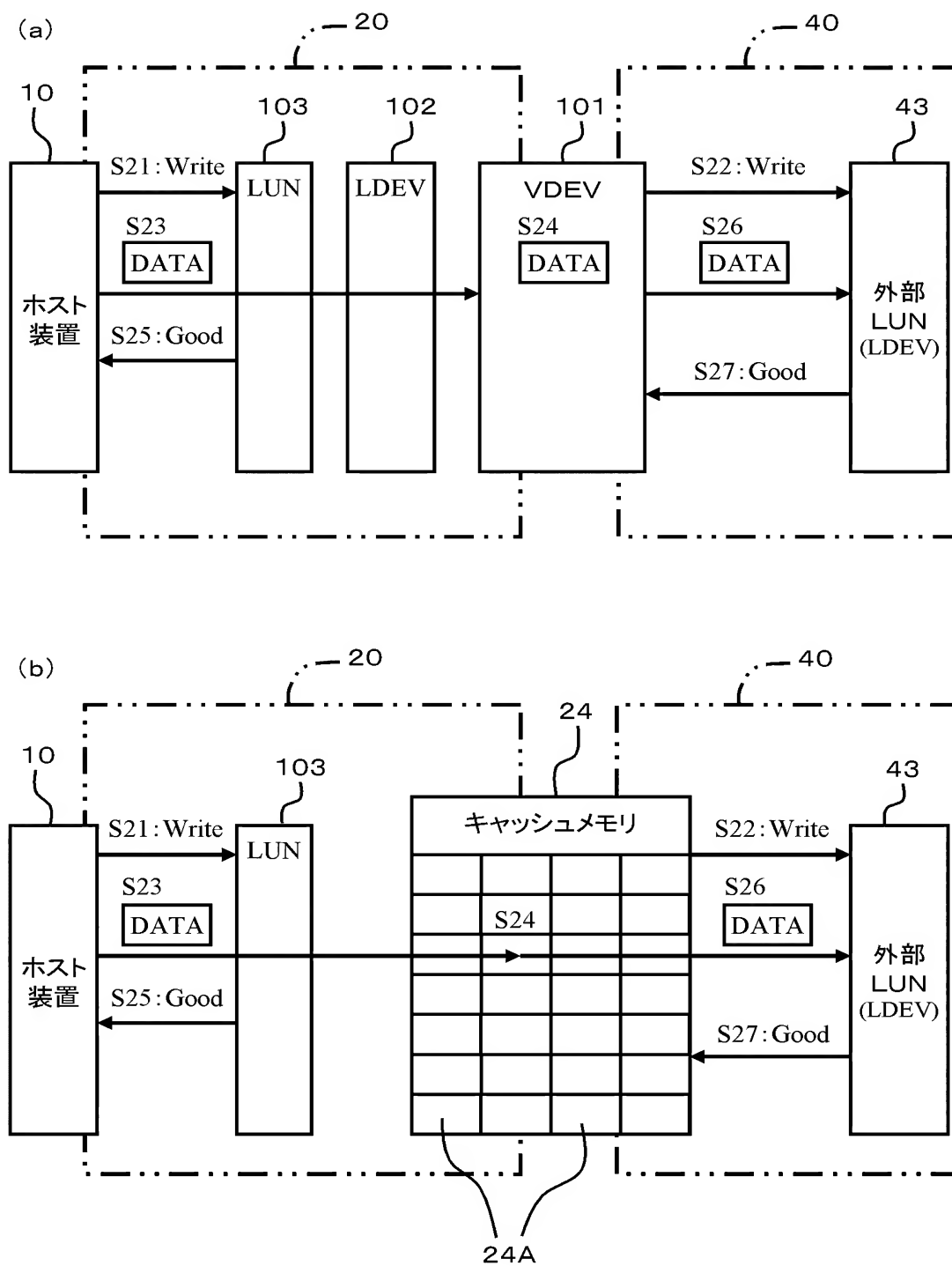
交換バス

Tm

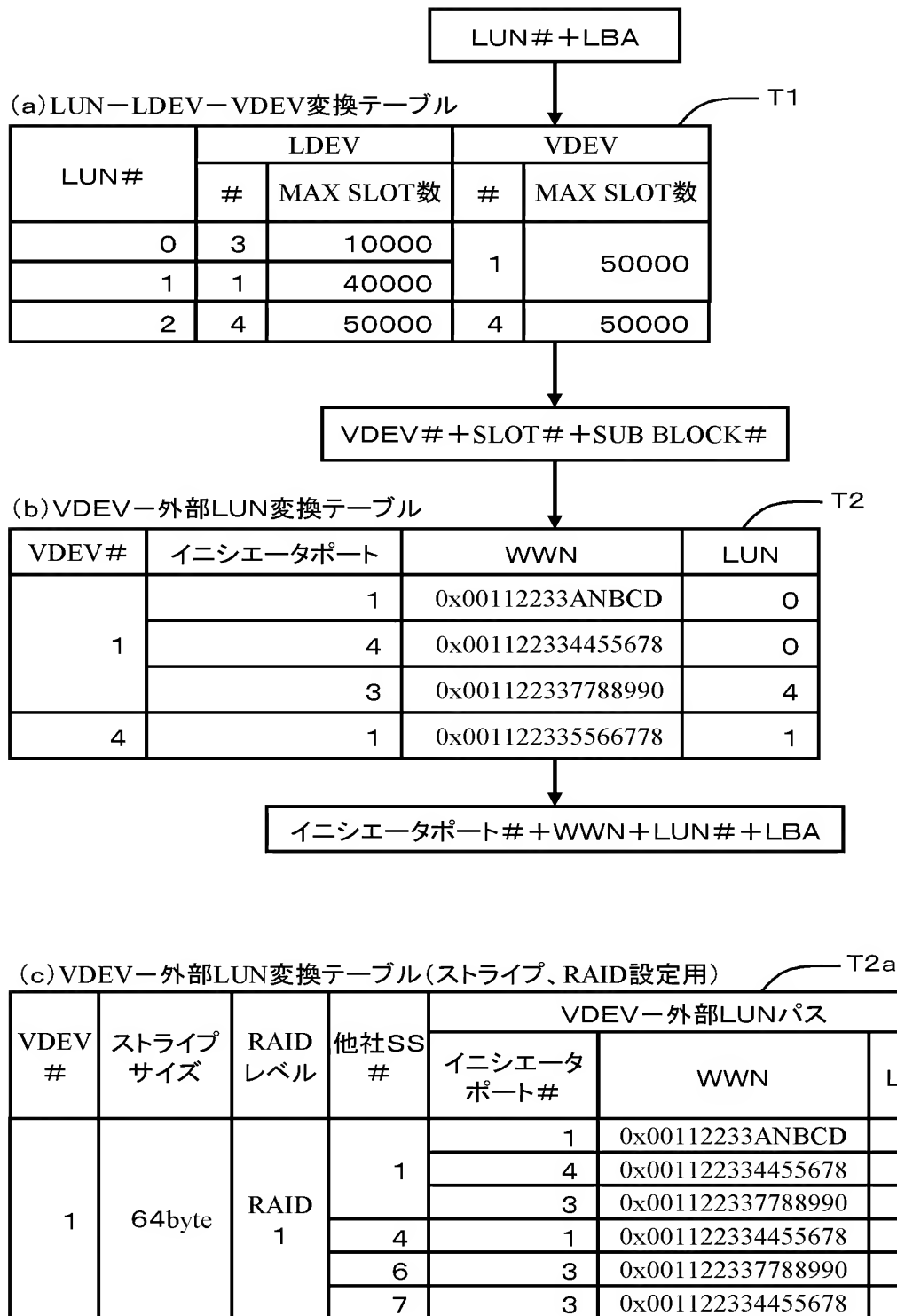
【図 5】



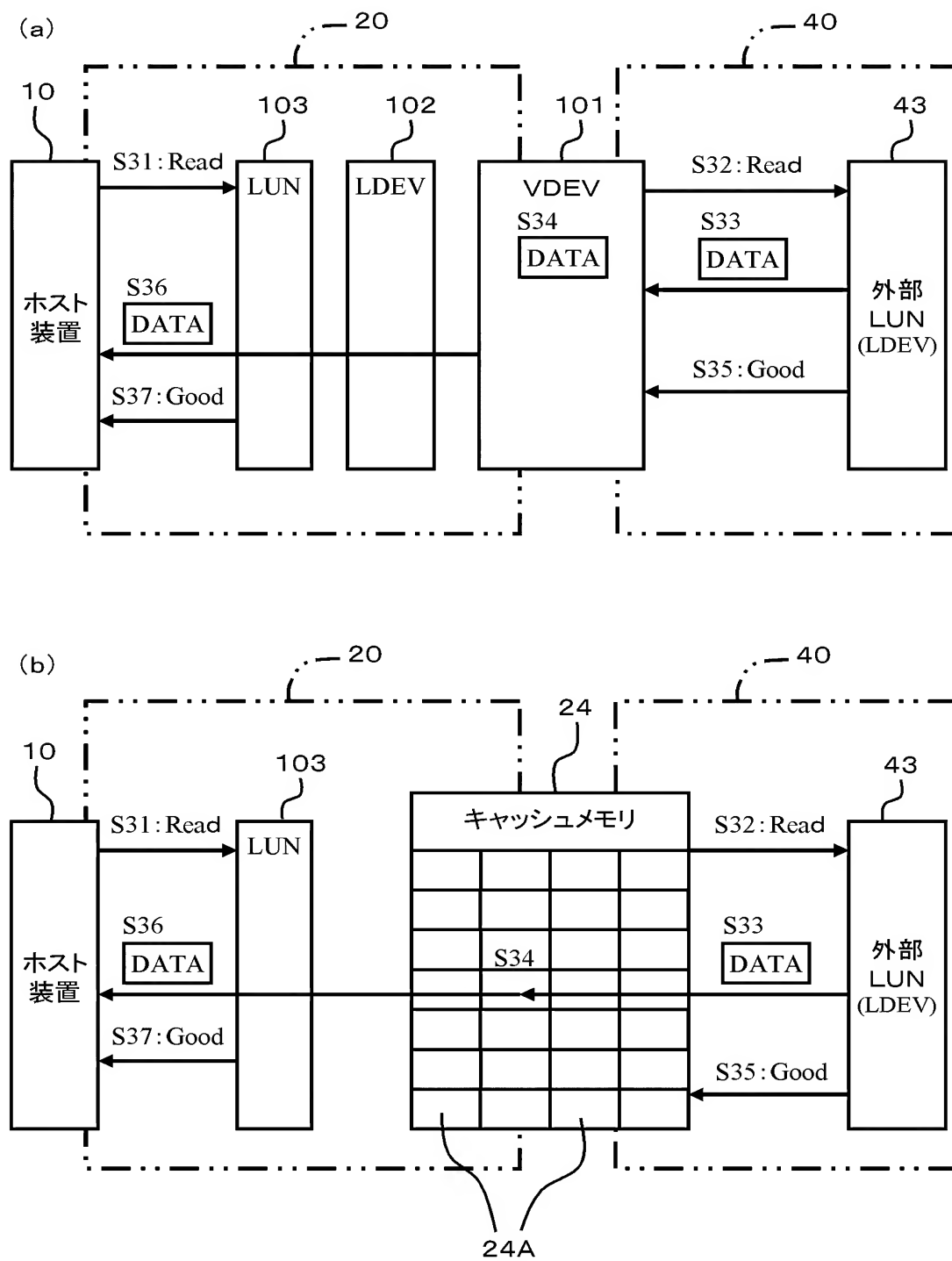
【図6】



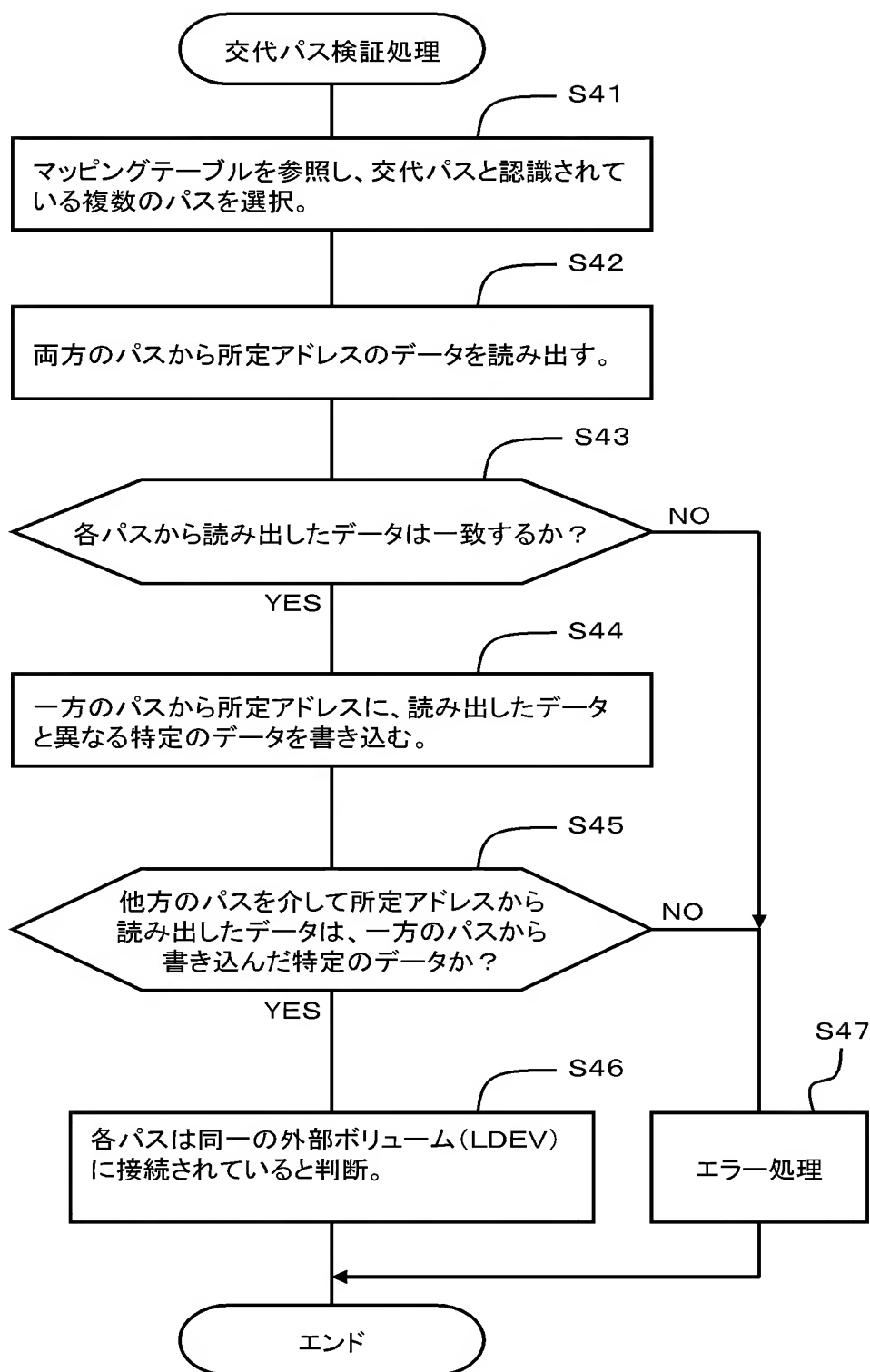
【図 7】



【図 8】



【図 9】



【書類名】 要約書

【要約】

【課題】 外部の記憶資源を内部の記憶資源として仮想化することにより、記憶資源を有効利用すること。

【解決手段】 第1の記憶制御装置20は、LUN103から接続されるLDEV（論理デバイス）102と、LDEV102の下位に接続されるVDEV（仮想デバイス）101との多層の記憶階層を有する。VDEV101の少なくとも1つは、外部の記憶制御装置40A～Dの有する記憶資源をマッピングすることにより構成されている。マッピングする際に、ストライプやRAID等の機能を追加することができる。外部の記憶資源を仮想的な内部の記憶資源として使用することで、通常の内蔵ボリュームに対して適用可能な各種機能（リモートコピー、可変ボリューム機能等）を、仮想的な内部ボリュームにも使用することができ、利用の自由度が高まる。

【選択図】 図2

【書類名】 手続補正書
【整理番号】 340300895
【あて先】 特許庁長官殿
【事件の表示】
【出願番号】 特願2003-337239
【補正をする者】
【識別番号】 000005108
【氏名又は名称】 株式会社日立製作所
【代理人】
【識別番号】 110000279
【氏名又は名称】 特許業務法人 ウィルフォート国際特許事務所
【代表者】 上村 輝之
【電話番号】 03-5296-7455
【連絡先】 担当は中村 猛

【手続補正1】

【補正対象書類名】 特許請求の範囲

【補正対象項目名】 全文

【補正方法】 変更

【補正の内容】

【書類名】 特許請求の範囲

【請求項1】

第1の記憶制御装置と第2の記憶制御装置とを通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶システムであって、

前記第1の記憶制御装置は、前記上位装置によりアクセスされる少なくとも1つ以上の論理ユニットと、この論理ユニットと少なくとも1つ以上の記憶デバイスとを接続するようにして設けられる少なくとも1つ以上の中間記憶階層とを備えており、

前記中間記憶階層は、前記記憶デバイス側に設けられる少なくとも1つ以上の第1の記憶階層と、この第1の記憶階層に接続され、かつ、前記論理ユニット側に設けられる少なくとも1つ以上の第2の記憶階層とを備えて構成されており、

前記第1の記憶階層が前記記憶デバイスに接続され、かつ、前記第2の記憶階層が前記論理ユニットに接続されることにより、前記第2の記憶制御装置が有する前記記憶デバイスを、前記中間記憶階層を介して前記論理ユニットに接続させる、記憶システム。

【請求項2】

上位装置及び第2の記憶制御装置と通信可能に接続され、前記上位装置からのリクエストに応じたデータ処理を行う記憶制御装置であって、

前記上位装置によりアクセスされる少なくとも1つ以上の論理ユニットと、前記論理ユニットと少なくとも1つ以上の記憶デバイスとを接続するようにして設けられる少なくとも1つ以上の中間記憶階層とを備え、

前記中間記憶階層は、前記記憶デバイス側に設けられる少なくとも1つ以上の第1の記憶階層と、この第1の記憶階層に接続され、かつ、前記論理ユニット側に設けられる少なくとも1つ以上の第2の記憶階層とを備えて構成されており、

前記第1の記憶階層が前記記憶デバイスに接続され、かつ、前記第2の記憶階層が前記論理ユニットに接続されることにより、前記第2の記憶制御装置が有する前記記憶デバイスを、前記中間記憶階層を介して前記論理ユニットに接続させる、記憶制御装置。

【請求項3】

前記上位装置からそれぞれ異なる複数の経路を介してアクセス可能な複数の論理ユニットを有し、

前記各論理ユニットは、前記中間記憶階層にそれぞれ接続されている請求項2に記載の記憶制御装置。

【請求項4】

前記第 2 の記憶制御装置が有する前記記憶デバイスへのパス情報を取得するパス情報取得手段を備え、

前記取得したパス情報が複数存在する場合は、前記各パス情報を同一の記憶デバイスへのパス情報として認識する請求項 2 に記載の記憶制御装置。

出願人履歴

000005108

19900831

新規登録

東京都千代田区神田駿河台4丁目6番地

株式会社日立製作所

000005108

20040908

住所変更

東京都千代田区丸の内一丁目6番6号

株式会社日立製作所